A General Analytical Framework for the Ising Model

with Applications to Inverse Ising Analysis of Purkinje Cell Neurodynamics

Phoenix Nénar Williams

Master of Science in Applied Mathematics
University of Colorado Denver

Thesis Defense
April 27, 2026

Dedication

"All we have to do is decide what to do with the time that is given to us."
— J.R.R. Tolkien

Acknowledgements: Mentorship and Research

Teachers and Mentors

Math teachers: Loreva Bromley, Larry Swanson, Jeff Garritano

Teachers/professors: Kate Nichols; Matthew Bolton, PhD

Mentors: Sister Kathleen Ann; Greg Sparks

Research Group

Advisor: Masoud Asadi-Zeydabadi, PhD

Research lead: Mazen Al-Borno, PhD

Graduate collaborators: Austin Long, Gunnar Enserro

Undergraduate collaborators: Sadiqah Al-Masyabi, Hanad Ali, Gabriel Adas

Data: Person Lab, Anschutz Medical Campus

Acknowledgements: Community and Family

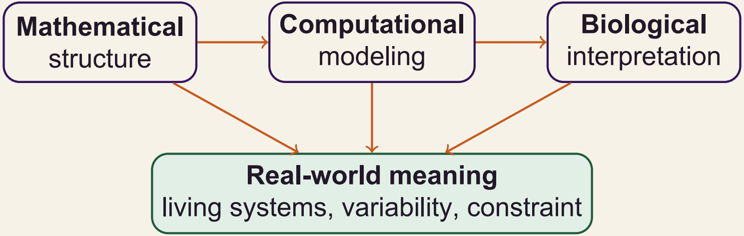
Community and Friends

Awesome Trans Masc Pastor BFF: Zoe
Oldest Friends: Madi and Robby
Trans Math Friends: Aubrey and River
Trans Daggerheart Party: Lexi and Dorian
Fellowship of the Boulder: Mostafa and Jessica
Other Trans Friends: Oliver, Jessie, Jack, and Darcy

Family

Partner: Jonce Palmer
Siblings: Zach and Kai
Doggos: Anigi, Brie, and Miso
Parents: Sara and Steven

Mathematics in Service of Living Systems



Long-term aim

Build tools that are both **computationally effective** and **biologically meaningful**.

Guiding Questions

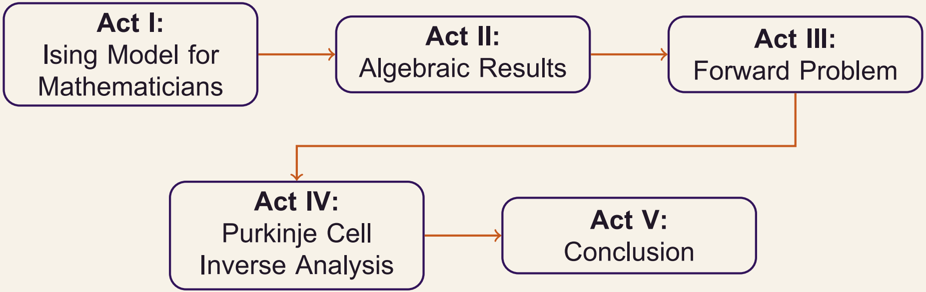
Mathematical Framing
How can the Ising model be written in language more natural to mathematicians and statisticians?

Forward Problem
What structure appears when the partition function is expanded algebraically?

Inverse Problem
Which computational choices most affect inverse Ising model fit?

Neural Application
How well do fitted models reproduce higher-order Purkinje cell activity?

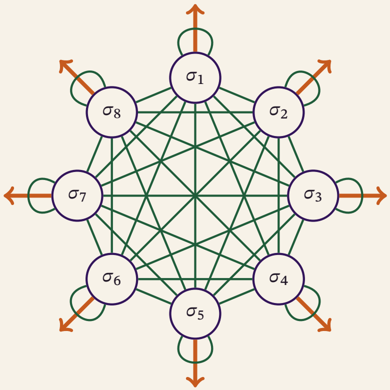
Roadmap


Act I:
Ising Model for Mathematicians
Act II:
Algebraic Results
Act III:
Forward Problem
Act IV:
Purkinje Cell Inverse Analysis
Act V:
Conclusion

Ising Model for Mathematicians

Act I

Core Objects



edges: $J_{ij} = J_{ji}$
loops: J_{ii}
arrows: h_i

- $\boldsymbol{\sigma} \in C_2^N$ (spin configuration)

  $\sigma_i = +1$: spin up
  $\sigma_i = -1$: spin down

- $\mathbf{J} \in \mathrm{Sym}_N(\mathbb{R})$ (pairwise interactions)

  $J_{ij} > 0$: ferromagnetic (edge present)
  $J_{ij} < 0$: antiferromagnetic (edge present)
  $J_{ij} = 0$: no edge / no interaction

- $\mathbf{h} \in \mathbb{R}^N$ (external field)

  $h_i > 0$: aligns with field
  $h_i < 0$: opposes field
  $h_i = 0$: no interaction

Energy and Probability

Hamiltonian (energy of a configuration)

$$\mathcal{E}(\boldsymbol{\sigma}) = -\boldsymbol{\sigma}^\top \mathbf{J}\boldsymbol{\sigma} - \mu\,\mathbf{h}^\top \boldsymbol{\sigma}$$

μ: magnetic moment

Boltzmann distribution

$$P_\beta(\boldsymbol{\sigma}) = \frac{\exp\left(-\beta\mathcal{E}(\boldsymbol{\sigma})\right)}{Z(\beta)} \quad \text{where} \quad Z(\beta) = \sum_{\boldsymbol{\sigma}' \in C_2^N} \exp\left(-\beta\mathcal{E}(\boldsymbol{\sigma}')\right)$$

$$\beta = \frac{1}{k_B T}, \quad |C_2^N| = 2^N$$

Mean Thermodynamic Quantities

Expectation-based formulation

$$E(\cdot) := E(\cdot \mid \mathbf{J}, \mathbf{h}, k_B, T)$$

Average energy

$$E(\mathcal{E}) = -\frac{\partial}{\partial \beta} \ln Z(\beta)$$

Magnetization

$$\overline{M} = E(\sigma)$$

Entropy

$$S = -k_B E(\ln P_\beta) = k_B \left(\beta E(\mathcal{E}) + \ln Z(\beta)\right)$$

Free energy

$$A = -k_B T \ln Z(\beta)$$

Heat capacity

$$C = \frac{\partial \mathcal{E}}{\partial T} + \overline{M} \mu \nabla \mathbf{h}$$

Susceptibility

$$\overline{\chi} = \frac{\overline{M}}{\mu E(\mathbf{h})} \quad \text{or} \quad \chi = \nabla_{\mathbf{h}} \overline{M} = \left(\frac{\partial \overline{M}_i}{\partial h_j}\right)_{i,j=1}^{N}$$

Thermodynamic Limit

As the system size grows, $N \to \infty$, thermodynamic quantities are typically normalized on a per-spin basis.

Magnetization

$$m = \frac{\overline{M}}{N}$$

Energy

$$u = \frac{E(\mathcal{E})}{N}$$

Entropy

$$s = \frac{S}{N}$$

Free Energy

$$a = \frac{A}{N}$$

Heat Capacity

$$c = \frac{C}{N}$$

Susceptibility

$$\chi = \frac{\overline{\chi}}{N}$$

Critical phenomena, such as spontaneous magnetization, appear below the Curie temperature T_c only under this limiting behavior.

Forward and Inverse Ising Problems

Forward problem

Given $(\mathbf{J}, \mathbf{h})$, analyze $P_\beta(\boldsymbol{\sigma})$

- Partition function
- Expectations
- High-probability states

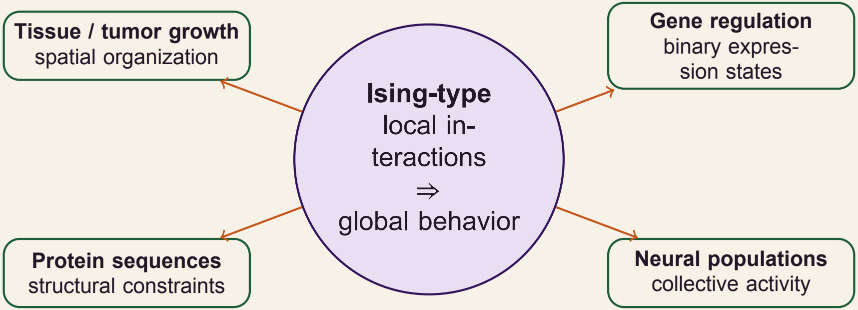
Inverse problem

Given data, estimate $(\mathbf{J}, \mathbf{h})$

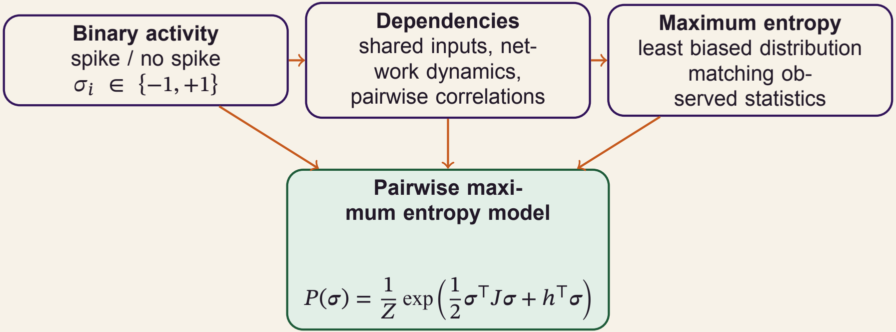
- Match empirical statistics
- Maximum likelihood / entropy
- Optimization problem

Both problems are computationally hard due to $Z(\beta)$.

Biological and Biomedical Applications



Why the Ising Model for Neural Data?

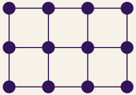


Mathematical Structure of the Ising Model

Act II

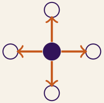
Computing the Partition Function - Physics View

Lattice geometry



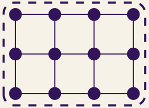
Dimension, geometry, and spatial arrangement.

Interaction neighborhood



Nearest-neighbor or other allowed interactions.

Boundary / manifold structure



Open, fixed, or periodic assumptions.

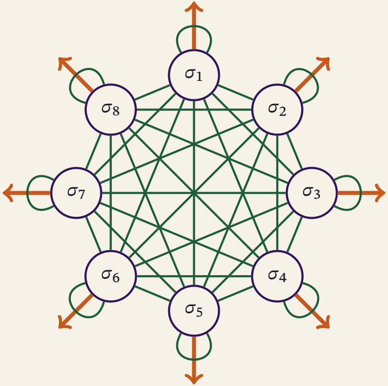
Sparsity pattern in $\mathbf{J}$

$$\mathbf{J} = \begin{pmatrix} 0 & * & 0 & 0 \\ * & 0 & * & 0 \\ 0 & * & 0 & * \\ 0 & 0 & * & 0 \end{pmatrix}$$

Many couplings are fixed to zero.

Transfer-matrix methods exploit structure in the lattice and local interactions.

Mathematical / Algebraic View - Quadratic Form Over C_2^N



Motivation

- Remove assumptions tied to geometry, dimension, and boundary conditions
- Separate intrinsic algebraic structure from modeling choices
- Enable analysis that applies to arbitrary interaction graphs

Benefits

- Results that generalize beyond lattice-based systems
- Clearer connection to combinatorics and linear algebra
- Alignment with graphical models and exponential families

Enumeration Scheme for C_2^N

Example: C_2^3

i	$i-1$	σ_i
1	000	$(+1,+1,+1)$
2	001	$(+1,+1,-1)$
3	010	$(+1,-1,+1)$
4	011	$(+1,-1,-1)$
5	100	$(-1,+1,+1)$
6	101	$(-1,+1,-1)$
7	110	$(-1,-1,+1)$
8	111	$(-1,-1,-1)$

Inductive description

1. Start with
$$\sigma_1 = (+1,\ldots,+1).$$
2. For $i \geq 2$, obtain σ_i by flipping the last coordinate with value $+1$ and all subsequent coordinates.

Binary formulation

$$\sigma_{i,j} = (-1)^{b_j(i-1)}$$

where $b_j(i-1)$ is the j^{th} bit of the binary representation of $i-1$.

Decomposing the Hamiltonian

Hamiltonian

$$\mathcal{E}(\boldsymbol{\sigma}) = -\Bigg(\underbrace{\boldsymbol{\sigma}^\top \mathbf{J}\boldsymbol{\sigma}}_{\mathcal{E}_J(\boldsymbol{\sigma})} + \underbrace{\mu\,\mathbf{h}^\top \boldsymbol{\sigma}}_{\mathcal{E}_h(\boldsymbol{\sigma})}\Bigg)$$

Interaction term

$$\mathcal{E}_J(\boldsymbol{\sigma}) = \boldsymbol{\sigma}^\top \mathbf{J}\boldsymbol{\sigma}$$

Field term

$$\mathcal{E}_h(\boldsymbol{\sigma}) = \mu\,\mathbf{h}^\top \boldsymbol{\sigma}$$

Simplifies algebraic analysis by separating interaction and field contributions.

Hadamard Multiplication

Interaction Term

$$\mathcal{E}_J(\boldsymbol{\sigma}) = \boldsymbol{\sigma}^\top \mathbf{J} \boldsymbol{\sigma}$$

rewrite as sum

$$\mathcal{E}_J(\boldsymbol{\sigma}) = \sum_{i,j} J_{ij} \sigma_i \sigma_j$$

$\sigma_i \sigma_j = (\sigma \sigma^T)_{ij}$

$$\mathcal{E}_J(\boldsymbol{\sigma}) = \sum_{i,j} (\mathbf{J} \odot \boldsymbol{\sigma}\boldsymbol{\sigma}^\top)_{ij}$$

$\sum A_{ij} = \mathbf{1}^T A \mathbf{1}$

$$\mathcal{E}_J(\boldsymbol{\sigma}) = \mathbf{1}^\top (\mathbf{J} \odot \boldsymbol{\sigma}\boldsymbol{\sigma}^\top) \mathbf{1}$$

Field Term

$$\mathcal{E}_h(\boldsymbol{\sigma}) = \mu\, \mathbf{h}^\top \boldsymbol{\sigma}$$

rewrite as sum

$$\mathcal{E}_h(\boldsymbol{\sigma}) = \mu \sum_i h_i \sigma_i$$

place terms on diagonal

$$\mathcal{E}_h(\boldsymbol{\sigma}) = \sum_{i,j} (\mu \operatorname{diag}(\mathbf{h}) \odot \operatorname{diag}(\boldsymbol{\sigma}))_{ij}$$

$\sum A_{ij} = \mathbf{1}^T A \mathbf{1}$

$$\mathcal{E}_h(\boldsymbol{\sigma}) = \mathbf{1}^\top (\mu \operatorname{diag}(\mathbf{h}) \odot \operatorname{diag}(\boldsymbol{\sigma})) \mathbf{1}$$

$$\mathcal{E}(\boldsymbol{\sigma}) = -\mathbf{1}^T \left(\mathbf{J} \odot \boldsymbol{\sigma}\boldsymbol{\sigma}^T + \mu \operatorname{diag}(\mathbf{h}) \odot \operatorname{diag}(\boldsymbol{\sigma}) \right) \mathbf{1}$$

Outer Products on C_2^N

General outer-product structure

$$
\begin{aligned}
(\sigma_i\sigma_i^T)_{j,k} &= \sigma_{i,j}\sigma_{i,k} \\
&= (-1)^{b_j(i-1)}(-1)^{b_k(i-1)} \\
&= (-1)^{b_j(i-1)+b_k(i-1)} \\
&= \begin{cases} -1 & b_j(i-1) \neq b_k(i-1) \\ 1 & b_j(i-1) = b_k(i-1) \end{cases} \\
&= \begin{cases} -1 & b_j(i-1) + b_k(i-1) = 1 \\ 1 & b_j(i-1) + b_k(i-1) \in \{0, 2\} \end{cases}
\end{aligned}
$$

Self-Connections

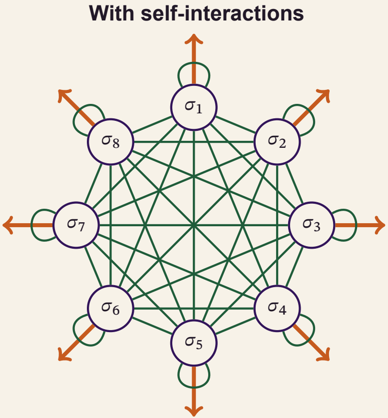
Diagonal Invariance to Spin Configuration:

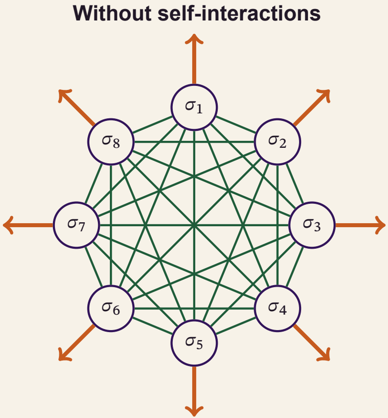
$$(\sigma_i\sigma_i^T)_{j,j} = (-1)^{2(b_j(i-1))} = (-1)^0 \text{ or } (-1)^2 = 1.$$

Self-connections in $\mathbf{J}$ do not affect the optimizing spin configuration.

$$J_{ii} = 0 \qquad \text{assumed going forward.}$$

Removing Self-Interactions





Complement Symmetry

$\sigma_{2^N-i} = -\sigma_i$

Binary Bitwise Complement:

$$(b_1(i-1),\dots,b_N(i-1)) \mapsto (1-b_1(i-1),\dots,1-b_N(i-1))$$

Integer Representation:

$$\begin{aligned}\sum_{j=1}^{N}(1-b_j(i-1))2^{N-j} &= \sum_{j=1}^{N}2^{N-j} - \sum_{j=1}^{N}b_j(i-1)2^{N-j}\\ &= ((2^N-1)-1)-(i-1)\\ &= 2^N-i-1.\end{aligned}$$

This integer matches spin configuration $(2^N-i-1)+1 = 2^N-i$, so:

$$\sigma_{2^N-i,j} = (-1)^{1-b_j(i-1)} = -(-1)^{b_j(i-1)} = -\sigma_{i,j}$$

$\sigma_{2^N-i}\sigma_{2^N-i}^T = \sigma_i\sigma_i^T$

Using

$$\sigma_{2^N-i} = -\sigma_i,$$

we obtain

$$\begin{aligned}(\sigma_{2^N-i}\sigma_{2^N-i}^T)_{j,k} &= \sigma_{2^N-i,j}\sigma_{2^N-i,k}\\ &= (-\sigma_{i,j})(-\sigma_{i,k})\\ &= \sigma_{i,j}\sigma_{i,k}\\ &= (\sigma_i\sigma_i^T)_{j,k}.\end{aligned}$$

How Many Negative Entries Can Occur?

Counting Negative Entries

Let

$$k = |\{j : \sigma_j = +1\}|, \qquad N - k = |\{j : \sigma_j = -1\}|.$$

Negative entries at: $\sigma_j \neq \sigma_k$.

Choices per ordering of $+1$ and -1: $k(N-k)$ each.

Thus,

$$f(k) = 2k(N-k).$$

Using Hamming weights $w_i = \sum_{j=1}^{N} b_j(i-1)$:

$$\begin{aligned} f(i) &= 2w_i(N - w_i) \\ &= 2\left(\sum_{j=1}^{N} b_j(i-1)\right)\left(N - \sum_{j=1}^{N} b_j(i-1)\right). \end{aligned}$$

Maximum Negative Entries

If N is even,

- Optimal Point: $k = N/2$
- Optimal Solution: $f(k) = N^2/2$

If N is odd,

- Optimal Point: $k = (N \pm 1)/2$
- Optimal Solution: $f(k) = (N^2 - 1)/2$

Therefore,

$$\max f(k) = 2\left\lfloor \frac{N^2}{4} \right\rfloor.$$

Multivariate Binomial Expansion

General form

For $x_1, \ldots, x_d, y_1, \ldots, y_d \in \mathbb{R}$ and integers $n_1, \ldots, n_d \geq 0$,

$$\prod_{i=1}^{d} (x_i + y_i)^{n_i} = \sum_{k_1=0}^{n_1} \cdots \sum_{k_d=0}^{n_d} \prod_{i=1}^{d} \binom{n_i}{k_i} x_i^{k_i} y_i^{n_i - k_i}.$$

Specialization: $n_i = 1$ for all $i \in \{1, \ldots, d\}$

$$\prod_{i=1}^{d} (x_i + y_i) = \sum_{k_1=0}^{1} \cdots \sum_{k_d=0}^{1} x_1^{k_1} y_1^{1-k_1} \cdots x_d^{k_d} y_d^{1-k_d} = \sum_{\mathbf{k} \in \{0,1\}^d} \prod_{i=1}^{d} x_i^{k_i} y_i^{1-k_i} = \sum_{S \subseteq \{1,\ldots,d\}} \left(\prod_{i \in S} x_i \right) \left(\prod_{i \notin S} y_i \right).$$

Character-Sum Cancellation

Let $c_1, \ldots, c_N \in \mathbb{Z}$. As k ranges over $\{1, \ldots, 2^N\}$,

$$k \longmapsto \mathbf{u} = (b_1(k-1), \ldots, b_N(k-1)) \in \{0,1\}^N.$$

Thus,

$$\sum_{k=1}^{2^N} (-1)^{\sum_{i=1}^N c_i b_i(k-1)} = \sum_{\mathbf{u}\in\{0,1\}^N} (-1)^{\sum_{i=1}^N c_i u_i}.$$

All c_i even

$$(-1)^{\sum c_i u_i} = 1$$

$$\implies \sum_{\mathbf{u}\in\{0,1\}^N} 1 = 2^N$$

Some c_r odd

Pair each $\mathbf{u}$ with $\mathbf{u} + e_r$:

$$(-1)^{\sum c_i u_i} + (-1)^{\sum c_i (u_i + e_{r,i})} = 0$$

All terms cancel.

$$\sum_{k=1}^{2^N} (-1)^{\sum_{i=1}^N c_i b_i(k-1)} = \begin{cases} 2^N, & c_i \equiv 0 \pmod 2 \text{ for all } i, \\ 0, & \text{otherwise.} \end{cases}$$

From Enumeration to Hyperbolic Form

$$Z(\beta) = \sum_{k=1}^{2^N} \exp\left(2\beta \sum_{i<j} J_{ij}\sigma_i^{(k)}\sigma_j^{(k)} + \beta\mu \sum_i h_i \sigma_i^{(k)}\right)$$

$\exp(\sum a_i) = \prod e^{a_i}$

$$Z(\beta) = \sum_{k=1}^{2^N} \left(\prod_{i<j} e^{2\beta J_{ij}\sigma_i^{(k)}\sigma_j^{(k)}}\right)\left(\prod_i e^{\beta\mu h_i \sigma_i^{(k)}}\right)$$

$e^{x\sigma} = \cosh(x) + \sigma\sinh(x)$

$$Z(\beta) = \sum_{k=1}^{2^N} \left(\prod_{i<j}[\cosh(2\beta J_{ij}) + \sigma_i^{(k)}\sigma_j^{(k)}\sinh(2\beta J_{ij})]\right) \cdot \left(\prod_i [\cosh(\beta\mu h_i) + \sigma_i^{(k)}\sinh(\beta\mu h_i)]\right)$$

Expanding by Edge and Vertex Subsets

Let $E_N = \{\{i,j\} : 1 \leq i < j \leq N\}$.

Choose $\sinh(2\beta J_{ij})$ for $\{i,j\} \in F \subseteq E_n$, and choose $\sinh(\beta\mu h_i)$ for $i \in S \subseteq \{1,\dots,N\}$

$$Z(\beta) = \sum_{k=1}^{2^N} \sum_{F\subseteq E_N} \sum_{S\subseteq\{1,\dots,N\}} \left(\prod_{\{i,j\}\in F} \sigma_i^{(k)}\sigma_j^{(k)} \sinh(2\beta J_{ij})\right)\left(\prod_{\{i,j\}\in E_N\setminus F} \cosh(2\beta J_{ij})\right) \cdot \left(\prod_{i\in S} \sigma_i^{(k)} \sinh(\beta\mu h_i)\right)\left(\prod_{i\notin S} \cosh(\beta\mu h_i)\right)$$

Factor spins, Rearrange Terms

$$Z(\beta) = \sum_{F\subseteq E_N} \sum_{S\subseteq\{1,\dots,N\}} \left(\prod_{\{i,j\}\in F} \sinh(2\beta J_{ij})\right)\left(\prod_{\{i,j\}\in E_N\setminus F} \cosh(2\beta J_{ij})\right) \cdot \left(\prod_{i\in S} \sinh(\beta\mu h_i)\right)\left(\prod_{i\notin S} \cosh(\beta\mu h_i)\right) \sum_{k=1}^{2^N} \left(\prod_{\{i,j\}\in F} \sigma_i^{(k)}\sigma_j^{(k)}\right)\left(\prod_{i\in S} \sigma_i^{(k)}\right).$$

Parity Determines Which Terms Survive

For fixed $F \subseteq E_N$,

$$\deg_F(i) = \#\{j : \{i,j\} \in F\}, \qquad \partial F = \{i : \deg_F(i) \text{ is odd}\}.$$

Collect spin powers

Each vertex i appears in the edge product exactly $\deg_F(i)$ times, so

$$\left(\prod_{\{i,j\}\in F} \sigma_i^{(k)}\sigma_j^{(k)}\right)\left(\prod_{i\in S} \sigma_i^{(k)}\right) = \prod_{i=1}^{N} \left(\sigma_i^{(k)}\right)^{\deg_F(i)+\mathbf{1}_{\{i\in S\}}}.$$

Using $\sigma_i^{(k)} = (-1)^{b_i(k-1)}$, the inner spin sum becomes

$$\sum_{k=1}^{2^N} (-1)^{\sum_{i=1}^{N}(\deg_F(i)+\mathbf{1}_{\{i\in S\}})b_i(k-1)}.$$

Apply cancellation

Given $b_i(k-1) = u_i$, we have

$$\sum_{\mathbf{u}\in\{0,1\}^N} (-1)^{\sum_{i=1}^{N}(\deg_F(i)+\mathbf{1}_{\{i\in S\}})u_i}.$$

By the cancellation lemma, this equals 2^N exactly when

$$\deg_F(i) + \mathbf{1}_{\{i\in S\}} \equiv 0 \pmod 2 \qquad \text{for all } i.$$

Therefore,

$$S = \partial F,$$

and all other choices of S cancel.

Partition Function by Edge Subsets

Exact expansion

$$Z(\beta) = 2^N \sum_{F \subseteq E_N} \left(\prod_{\{i,j\} \in F} \sinh(2\beta J_{ij}) \right) \left(\prod_{\{i,j\} \in E_N \setminus F} \cosh(2\beta J_{ij}) \right) \cdot \left(\prod_{i \in \partial F} \sinh(\beta \mu h_i) \right) \left(\prod_{i \notin \partial F} \cosh(\beta \mu h_i) \right)$$

Advantages

- Fully analytical (elementary functions)
- Independent of lattice structure
- Encodes interactions via graph structure
- Systematic cancellation eliminates spin dependence

Limitations

- Still exponential in $|E_N| = \binom{N}{2}$
- Not optimized for direct computation
- Best suited for analysis, not simulation

What This Representation Enables

Thermodynamic quantities

- Derived from $Z(\beta)$ via logarithmic derivatives
- Parity constraints extend to expectations and correlations
- Suggests graph-based formulas for observables
- Full development deferred to future work

Model generality

- Does not fundamentally require symmetry of $\mathbf{J}$
- Extends to asymmetric / directional interactions
- Independent of lattice dimension or geometry
- Applicable to non-lattice and dense graphs
- Connection to layered motor cortex data

From summation over spin configurations $\longrightarrow$ structured sums over edge subsets and boundaries.

The Forward Ising Problem

Act III

Optimization Formulation and Challenges

Optimization formulations

Probability formulation

$$\begin{array}{ll} \text{maximize} & P_\beta(\boldsymbol{\sigma}) \\ \text{subject to} & \boldsymbol{\sigma} \in C_2^N \end{array}$$

$\Downarrow$

Energy formulation

$$\begin{array}{ll} \text{minimize} & -\boldsymbol{\sigma}^T \mathbf{J} \boldsymbol{\sigma} - \mu \boldsymbol{\sigma}^T \mathbf{h} \\ \text{subject to} & \boldsymbol{\sigma} \in C_2^N \end{array}$$

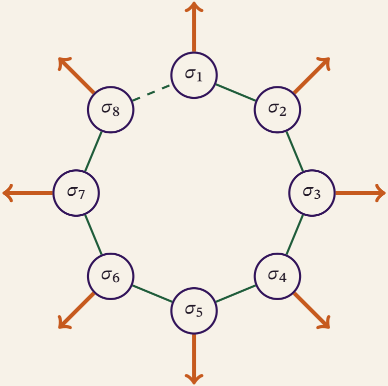
Computational challenges

- Discrete, non-convex domain C_2^N
- No differentiability → no gradient methods
- No guarantees on $\mathbf{J}$ (invertibility / definiteness)
- Equivalent to QUBO (NP-hard)
- Exhaustive search: $\mathcal{O}(N2^N)$

Exact solutions do not scale with system size.

One-Dimensional Ising Model

Nearest-neighbor structure (dashed = cyclic boundary)



Nearest-neighbor Hamiltonian

$$\mathcal{E}(\sigma) = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - h \sum_{i=1}^{N} \sigma_i$$

For cyclic boundaries,

$$\sigma_{N+1} = \sigma_1.$$

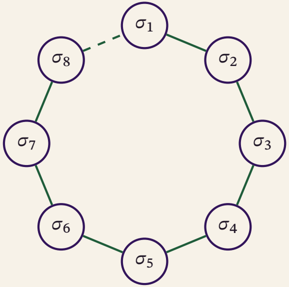
Transfer matrix

$$T = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta h} \end{pmatrix}$$

$$\lambda_{\pm} = e^{\beta J} \left[\cosh(\beta h) \pm \left(\sinh^2(\beta h) + e^{-4\beta J} \right)^{1/2} \right]$$

One-Dimensional Model with No External Field

Nearest-neighbor structure ($h = 0$, dashed = cyclic boundary)



Transfer-matrix simplification

$$\lambda_+ = 2\cosh(\beta J), \qquad \lambda_- = 2\sinh(\beta J)$$

F-sum interpretation ($h = 0$)

$$\sinh(0) = 0 \ \Rightarrow\ \partial F = \varnothing$$

- **Path:** only $F = \varnothing$
- **Cycle:** $F = \varnothing$ and $F = E$

Only even subgraphs survive → exact agreement with transfer-matrix formulas.

1D Thermodynamic Consequences

Thermodynamic limit

$$Z_N \sim \lambda_+^N$$

Free Energy

$$A = -\lim_{N\to\infty} \frac{1}{\beta N} \log Z_N = -\frac{1}{\beta} \log \lambda_+$$
$$h = 0 \implies A = -\frac{1}{\beta} \log(2\cosh(\beta J)).$$

Magnetization

$$m = \sinh(\beta h)\left[\sinh^2(\beta h) + e^{-4\beta J}\right]^{-1/2}$$
$$h = 0 \implies m = 0.$$

Energy and heat capacity for $h = 0$

$$u = -J\tanh(\beta J)$$
$$C = (\beta J)^2\left(1 - \tanh^2(\beta J)\right)$$

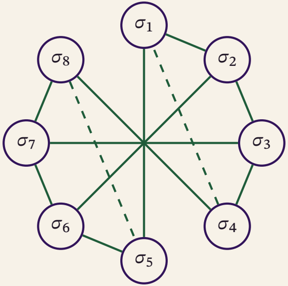
Correlations

$$\langle \sigma_i \sigma_{i+r} \rangle = \left(\frac{\lambda_-}{\lambda_+}\right)^r \sim e^{-r/\xi}$$
$$\xi^{-1} = -\log\left|\frac{\lambda_-}{\lambda_+}\right|.$$

Since $|\lambda_-/\lambda_+| < 1$ at finite temperature, correlations decay exponentially and no finite-temperature phase transition occurs in 1D.

Two-Dimensional Ising Model

Square-lattice nearest neighbors



Solid edges show nearest-neighbor interactions. Dashed edges indicate periodic/toroidal boundary identifications.

Zero-field Hamiltonian

$$\mathcal{E}(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

- First standard setting with a true phase transition
- Exact thermodynamic-limit solution due to Onsager
- Critical temperature:

$$\sinh(2\beta_c J) = 1 \quad \Longleftrightarrow \quad k_B T_c = \frac{2J}{\log(1+\sqrt{2})}$$

Consistency with the F-Sum (2D, $h = 0$)

Zero-field simplification

$$\sinh(0) = 0 \quad \Rightarrow \quad \partial F = \varnothing$$

Constraint

$$\deg_F(i) \equiv 0 \pmod 2$$

Interpretation

Every vertex has even degree

Allowed subgraphs are unions of closed loops.

Resulting expansion

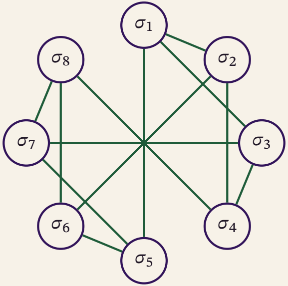
$$Z(\beta) = 2^{|V|} \left(\prod_{e \in E} \cosh(\beta J) \right) \sum_{\substack{F \subseteq E \\ \partial F = \varnothing}} \prod_{e \in F} \tanh(\beta J)$$

Key consequence

- Sum over even subgraphs
- Equivalent to loop (graphical) expansion
- Matches classical high-temperature expansion

Higher-Dimensional Ising Models

3D nearest-neighbor structure



Edges represent nearest-neighbor couplings in a 3D interaction structure.

Distinct 3D periodic (toroidal) boundary identifications require a larger lattice; this diagram preserves nearest-neighbor structure within a fixed node set.

What changes in higher dimensions?

- No general closed-form analytical solution is known for $d \geq 3$
- Phase transitions still occur for $d \geq 2$
- Thermodynamic quantities are studied through approximation and simulation
- Monte Carlo methods become central computational tools

F-sum consistency

$$Z(\beta) = \sum_{F \subseteq E} \text{weighted subgraph contributions}$$

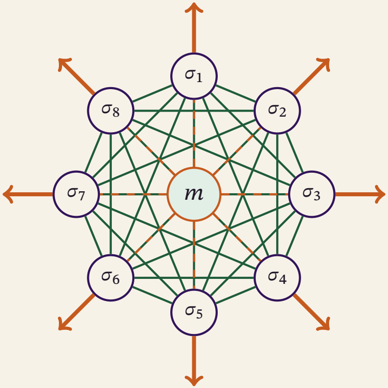
For $h = 0$,

$$\partial F = \varnothing$$

restricts the sum to even subgraphs.

Mean Field Theory

Dense interaction → average effect



Each spin interacts with the average magnetization m.

Self-consistency equation

$$m = \tanh\left(\beta(Jm + h)\right)$$

Interpretation

- Replace full interaction structure with a single scalar m
- Reduces high-dimensional problem to 1D optimization
- Captures phase transition at

$$k_B T_c = J$$

Limitations

- Neglects correlations and fluctuations
- Inaccurate critical behavior in low dimensions

Numerical Methods for the Ising Model

Why numerical methods?

$$|C_2^N| = 2^N$$

Direct computation of the partition function is infeasible for moderate N.

Instead, estimate expectations using samples from the Boltzmann distribution:

$$P(\boldsymbol{\sigma}) = \frac{e^{-\beta \mathcal{E}(\sigma)}}{Z(\beta)}.$$

Observable quantities (e.g., magnetization, energy, correlations) are approximated via sample averages.

Metropolis–Hastings algorithm

1. Initialize $\boldsymbol{\sigma}^{(0)}$
2. For each iteration:
   2.1 Select $i \in \{1, \dots, N\}$
   2.2 Propose flip: $\sigma_i \to -\sigma_i$
   2.3 Compute

   $$\Delta\mathcal{E} = \mathcal{E}(\boldsymbol{\sigma}') - \mathcal{E}(\boldsymbol{\sigma})$$

   2.4 Accept with probability

   $$\min\left(1, e^{-\beta\Delta\mathcal{E}}\right)$$

3. Repeat step 2. After a burn-in period, samples converge the Boltzmann distribution.

Improving Sampling Efficiency

Limitations of Metropolis–Hastings

- Local single-spin updates
- Slow mixing near T_c
- Long-range correlations slow convergence
- Complex energy landscapes can trap chains locally

Parallel tempering

- Simulate at multiple temperatures
- Swap configurations between replicas
- High-temperature replicas escape local minima
- Low-temperature replicas preserve target structure

Cluster methods

- Replace local flips with nonlocal updates
- Reduce critical slowing down

Swendsen–Wang

- Build clusters of aligned spins
- Flip clusters simultaneously

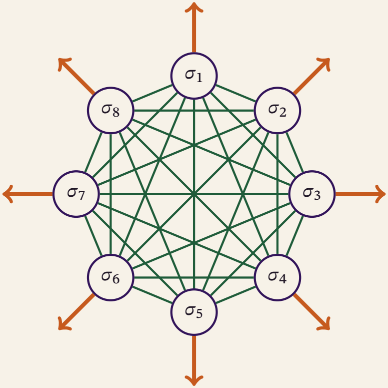
Wolff

- Grow one cluster per iteration
- Flip the selected cluster
- Often lower overhead than full-cluster updates

Inverse Ising Analysis of Purkinje Cell Neurodynamics

Act IV

Inverse Ising Problem



Fully connected model with external field.

Inverse Ising formulation

Given neural samples $\{\boldsymbol{\sigma}^{(m)}\}_{m=1}^{M} \subset C_2^N$, estimate $(\mathbf{J}, \mathbf{h})$ so that

$$\langle \sigma_i \rangle_{\text{model}} \approx \langle \sigma_i \rangle_{\text{data}}, \qquad \langle \sigma_i \sigma_j \rangle_{\text{model}} \approx \langle \sigma_i \sigma_j \rangle_{\text{data}}.$$

Optimization problem

$$\text{minimize } -\ell(\mathbf{J}, \mathbf{h}) = -\frac{1}{M} \sum_{m=1}^{M} \beta \left(\frac{1}{2} \boldsymbol{\sigma}^{(m)\top} \mathbf{J} \boldsymbol{\sigma}^{(m)} + \mathbf{h}^\top \boldsymbol{\sigma}^{(m)} \right) + \log Z(\mathbf{J}, \mathbf{h})$$

subject to $\mathbf{J} \in \text{Sym}_N(\mathbb{R})$

$\mathbf{h} \in \mathbb{R}^N$

Challenges and Computational Approach

Why optimization is hard

Nonlinear coupling via Z:

$$\nabla \ell(\mathbf{J}, \mathbf{h}) = E_{\text{data}}[\cdot] - E_{\text{model}}[\cdot]$$

- Model expectations depend on $Z(\mathbf{J}, \mathbf{h})$
- Coupling across all 2^N configurations
- Convex objective, but intractable in practice

Modeling considerations

- Fully connected graph ($\mathbf{J}$ dense)
- Pairwise interactions (no higher-order terms)
- Structural constraints / regularization optional

Approximate methods

Pseudolikelihood

- Factorize likelihood into conditionals
- Reduces to logistic regression

Monte Carlo

- Estimate $E_{\text{model}}[\cdot]$ via sampling
- Scales better but computationally intensive

Mean-field

- Replace interactions with average effect
- Fast, but ignores correlations

Approximation replaces exact evaluation of Z with tractable surrogates.

Inverse Ising Solvers (`coniii`)

Enumeration

- Exact evaluation of likelihood over all configurations
- Highest accuracy (ground truth benchmark)
- $\mathcal{O}(K^2N^2)$ $\quad$ ($K = 2^N$ configurations)

Monte Carlo Histogram (MCH)

- Sampling-based estimation of likelihood gradients
- Asymptotically exact (variance from sampling)
- $\mathcal{O}(TnN^2)$ $\quad$ (T iterations, n samples)

Adaptive Cluster Expansion (ACE)

- Expands model over subsets of variables (clusters)
- High accuracy for sparse/structured interactions
- $\mathcal{O}(\binom{N}{n}2^n)$ $\quad$ (n cluster size)

Pseudolikelihood

- Factorizes likelihood into conditional distributions
- Fast but biased (misses global correlations)
- $\mathcal{O}(R(N^2 + N^3))$ $\quad$ (R iterations)

Minimum Probability Flow (MPF)

- Minimizes probability flow without computing Z
- Good accuracy with sufficient data coverage
- $\mathcal{O}(RGN^2)$ $\quad$ (R iterations, G neighbors)

Regularized Mean Field (RMF)

- Covariance-based approximation (mean-field)
- Fast but lowest accuracy (ignores correlations)
- $\mathcal{O}(N^3)$

Expected Solver Performance (Theory)

Accuracy (theoretical)

Enumerate $>$ MCH $\approx$ ACE $>$ MPF $>$ Pseudo

- Exact likelihood recovers full distribution
- Sampling methods converge with sufficient samples
- Approximations lose higher-order dependencies

Bias characteristics

- **Pseudo:** underestimates correlations
- **MPF:** depends on chosen transition graph
- **MCH:** variance from finite sampling
- **ACE:** truncation error from cluster selection

Computational cost

Enumerate $\gg$ ACE $\gtrsim$ MCH $\gg$ MPF $\approx$ Pseudo

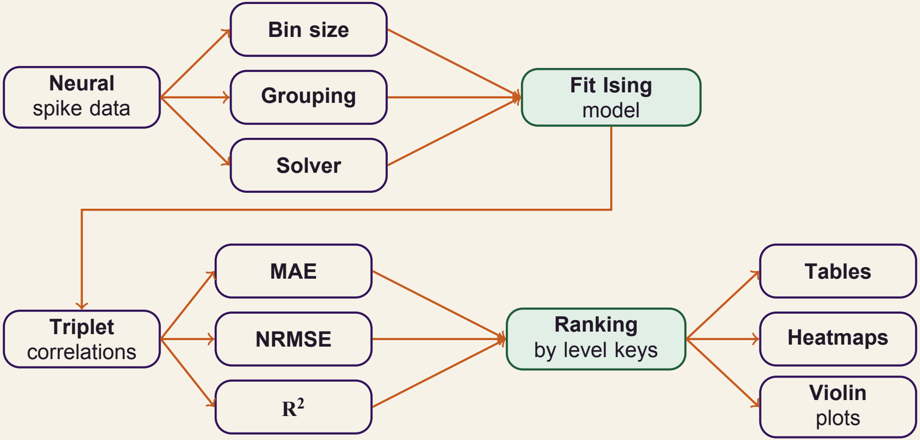
- Exponential vs polynomial scaling regimes
- Sampling cost dominates MCH runtime
- MPF / Pseudo scale to large N

Interpretation

Accuracy $\longleftrightarrow$ Scalability

- Left: exact but intractable
- Right: scalable but biased
- Middle: balance (MCH, ACE)

Factorial Search Pipeline



Person Purkinje Cell Data and Search Setup

Original file hierarchy

Mouse → Reach category → Session file → Stim condition → Reach

- 6 mice
- 3 stimulus states per session
- Reach status: all / fail / success

Levels

1. Overall
2. Stimulus State
3. Reach Status
4. Subject
5. Session/Date
6. Reach ID

Person-data search settings

- Bin sizes: $1, 5, 10, 15, 20, 25$ ms
- Grouping: by reach / by session
- Solvers: MPF, MCH, pseudo, ACE, Enumeration $N \leq 15$

Preprocessing

$$\sigma_i = \begin{cases} +1, & \text{spike in window} \\ -1, & \text{no spike in window} \end{cases}$$

ACE feasibility note

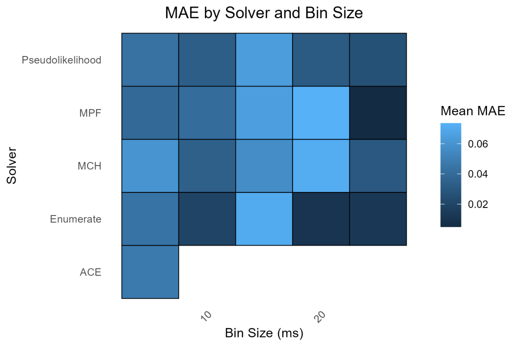
ACE succeeded at small bin sizes, but failed at larger bin sizes due to exceeding repeated runtime limits.

Triplet-Correlation MAE: Aggregation Table

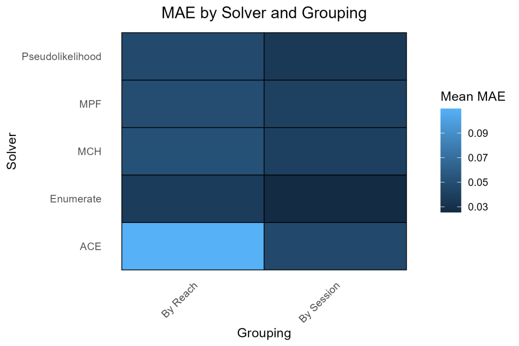
Level Type	Solver	Grouping	Bin Size (ms)	Mean MAE	SD	Count
Overall	MCH	By Session	10	0.06109		1
Stim	MCH	By Session	10	0.04853		1
	Enumerate	By Session	10	0.05888		1
	Pseudolikelihood	By Session	10	0.06331		1
Reach Status	Enumerate	By Session	10	0.04848		1
	Enumerate	By Session	5	0.05378		1
	Pseudolikelihood	By Session	10	0.05441	0.01118	4
Subject	Enumerate	By Session	10	0.01090		1
	Enumerate	By Session	15	0.02551	0.001485	2
	Enumerate	By Session	5	0.03738	0.01201	2
Session	Enumerate	By Session	20	0.0002654		1
	Enumerate	By Session	10	0.02026	0.02868	5
	Pseudolikelihood	By Session	20	0.02088	0.01557	2
Reach	Enumerate	By Session	20	0.0002654	0	15
	MCH	By Reach	20	0.002696	0	2
	MPF	By Reach	25	0.004988		1

Table: Top 3 best-performing configurations by mean MAE within each level type.

Triplet-Correlation MAE: Heatmaps

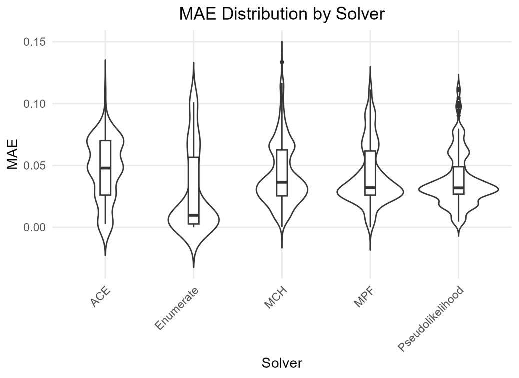


Solver × bin size

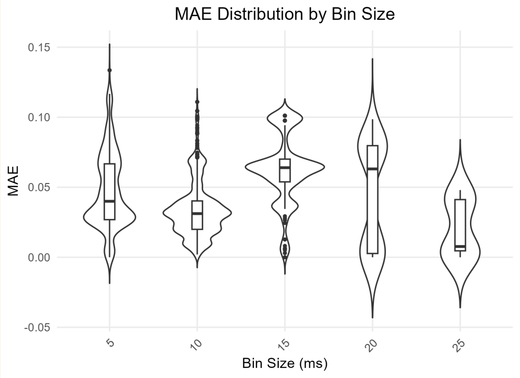


Solver × grouping

Triplet-Correlation MAE: Violin Plots



Distribution of MAE of triplet correlations by solver.



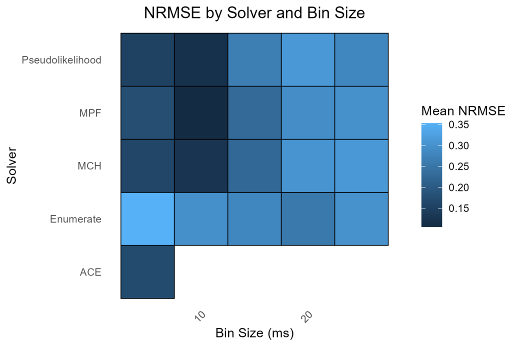
Distribution of MAE of triplet correlations by bin size.

Triplet-Correlation NRMSE: Aggregation Table

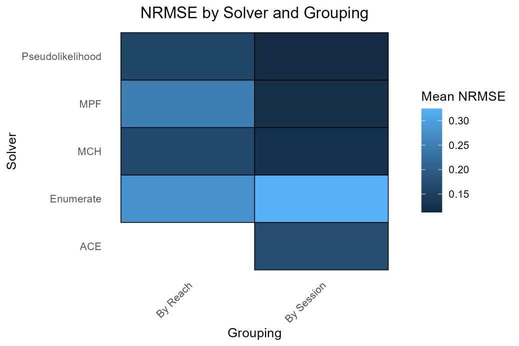
Level Type	Solver	Grouping	Bin Size (ms)	Mean NRMSE	SD	Count
Overall	Pseudolikelihood	By Reach	10	0.2293		1
Stim	MPF	By Session	10	0.1706		1
	Pseudolikelihood	By Reach	10	0.2257		1
	MCH	By Reach	10	0.2751		1
Reach Status	MCH	By Session	10	0.1422	0.04993	2
	MPF	By Session	10	0.1548	0.02798	3
	Pseudolikelihood	By Reach	10	0.2338	0.01034	2
Subject	Pseudolikelihood	By Reach	10	0.1123	0.002823	5
	MCH	By Session	10	0.1305	0.04047	17
	Pseudolikelihood	By Session	10	0.1326	0.05884	7
Session	Pseudolikelihood	By Session	10	0.09386	0.05261	18
	MPF	By Session	10	0.1062	0.07530	30
	MCH	By Session	10	0.1081	0.05723	35
Reach	Pseudolikelihood	By Session	10	0.09184	0.05245	315
	MPF	By Session	10	0.09287	0.06657	281
	MCH	By Session	10	0.1060	0.05400	425

Table: Top 3 best-performing configurations by mean NRMSE within each level type.

Triplet-Correlation NRMSE: Heatmaps

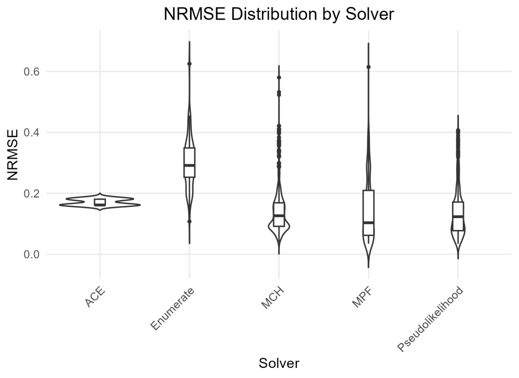


Solver × bin size

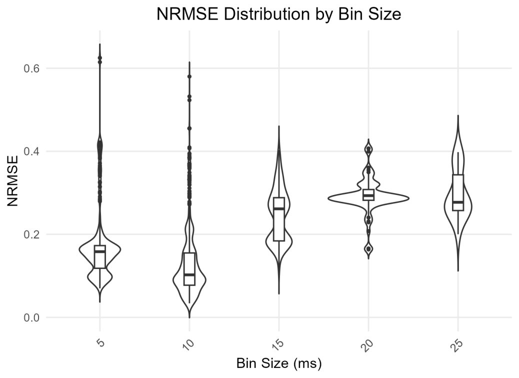


Solver × grouping

Triplet-Correlation NRMSE: Violin Plots



Distribution of NRMSE of triplet correlations by solver.



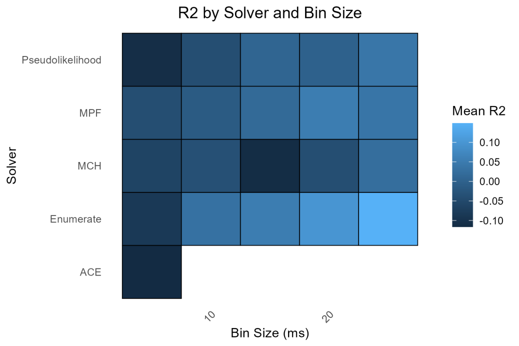
Distribution of NRMSE of triplet correlations by bin size.

Triplet-Correlation R^2: Aggregation Table

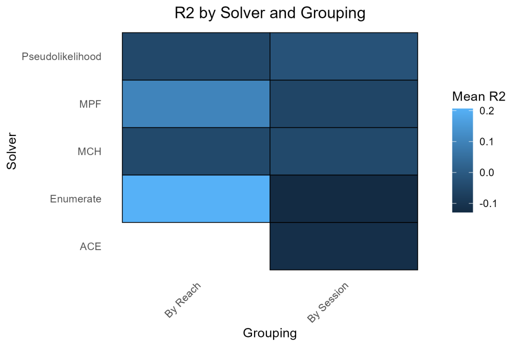
Level Type	Solver	Grouping	Bin Size (ms)	Mean R^2	SD	Count
Overall	MCH	By Reach	15	-1.375		1
Stim	MPF	By Session	10	-0.4857		1
	Pseudolikelihood	By Reach	10	-1.493		1
	MCH	By Reach	15	-1.747		1
Reach Status	Enumerate	By Session	10	0.3061	0.5502	2
	MPF	By Session	10	-0.4621		1
	MCH	By Session	10	-0.5547		1
Subject	Enumerate	By Reach	15	0.6125		1
	Enumerate	By Session	10	0.02699	0.3642	6
	Pseudolikelihood	By Reach	15	-0.02102		1
Session	Pseudolikelihood	By Session	15	0.2931	0.5521	3
	Pseudolikelihood	By Session	20	0.02037		1
	Enumerate	By Session	20	0.01907		1
Reach	Enumerate	By Reach	10	0.2589	0.2811	30
	MPF	By Reach	10	0.2030	0.2105	16
	Enumerate	By Reach	5	0.1880	0.2564	14

Table: Top 3 best-performing configurations by mean R^2 within each level type.

Triplet-Correlation R^2: Heatmaps

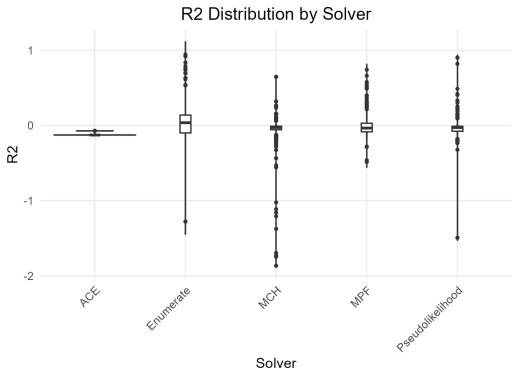


Solver × bin size

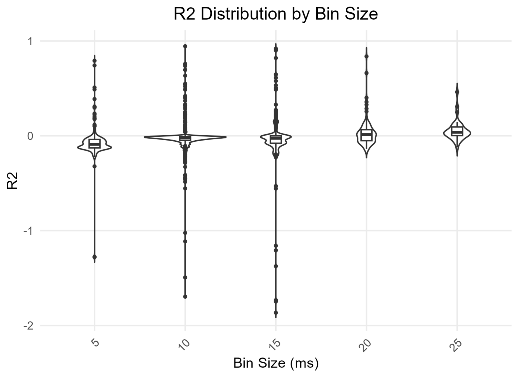


Solver × grouping

Triplet-Correlation R^2: Violin Plots



Distribution of R^2 of triplet correlations by solver.



Distribution of R^2 of triplet correlations by bin size.

Interpreting Performance Metrics

Error-based metrics (MAE, NRMSE)

- Consistent conclusions across metrics
- Best performance:
  - MPF, Pseudolikelihood
  - Bin sizes ≈ 10 ms
  - Grouping by reach
- Capture accurate magnitudes of correlations

Variance-based metric (R^2)

- Often near zero or negative
- Sensitive to low variance in data
- Enumeration performs better when variability exists

Key distinction

- MAE / NRMSE:
  - Minimize absolute error
  - Favor stable, low-variance reconstructions
- R^2:
  - Measures variance explained
  - Rewards preserving spread of data

No single metric fully captures model performance.

Solver choice determines how variability is represented.

Theory vs Computational Performance

Theoretical expectations

- Accuracy:

  Enumerate > MCH/ACE > MPF > Pseudo

- Cost:

  Enumerate $\gg$ ACE/MCH $\gg$ MPF/Pseudo

Observed behavior

- Enumeration: best but limited to $N \lesssim 15$
- ACE: infeasible (runtime explosion)
- MPF: slower than expected
- Pseudolikelihood + MCH: most stable

Data-dependent effects

- MPF sensitive to sample size + structure
- ACE fails due to lack of sparsity
- MCH benefits from practical sampling efficiency

Grouping trade-off

- *By reach*: faster per fit
- *By session*: fewer total fits

Solver performance depends on data structure, not just theory.

Conclusion

Act V

Summary of Contributions

Theoretical contributions

- Algebraic formulation over C_2^N
- Partition function via combinatorial expansion
- Reduced dependence on lattice assumptions

Computational contributions

- Factorial search over solvers, binning, grouping
- Multi-metric evaluation (MAE, NRMSE, R^2)
- Empirical validation on neural spike data

Key findings

- MPF and Pseudolikelihood minimize reconstruction error
- Enumeration preserves variability when feasible
- R^2 unstable in low-variance regimes

Solver performance depends on both algorithm design and data structure.

Key Insights

For Ising model analysis

- Unified algebraic + combinatorial framework
- Generalizes beyond lattice-based systems
- Connects forward and inverse problems

For inference

- No single metric is sufficient
- Trade-off: accuracy vs variability
- Theory alone does not predict performance

For neural modeling

- Pairwise models capture key structure
- Preprocessing (binning, grouping) is critical
- Solver choice shapes inferred dynamics

Accurate modeling requires integrating theory, computation, and data structure.

Future Directions

Theoretical

- Closed-form / asymptotic partition results
- Bayesian inverse Ising formulations
- Extensions to asymmetric interactions

Computational

- Scalable solvers for high-dimensional data
- Algorithms tailored to neural structure
- Improved handling of temporal dependence

Applications

- Additional datasets (motor + visual cortex)
- Cross-region validation

Interpretation

- Link $\mathbf{J}, \mathbf{h}$ to behavior
- Study criticality and scaling
- Integrate neural + kinematic data

Questions?

1D Partition Function with External Field

	Transfer Matrix	Edge-Subset / Algebraic
Cyclic	$Z_N^{(c)} = \lambda_+^N + \lambda_-^N = \lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right]$	$E = \{\{1,2\}, \dots, \{N,1\}\}$ $F =$ intervals or full cycles $\partial F =$ interval endpoints Each component gives two boundary vertices, except full cycles.
Free	$Z_N^{(f)} = \lambda_+^{N-1}\alpha_+ + \lambda_-^{N-1}\alpha_-$ $\alpha_\pm = \cosh(\beta h) \pm \frac{\sinh^2(\beta h) + e^{-2\beta J}}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}$	$E = \{\{1,2\}, \dots, \{N-1,N\}\}$ $F =$ disjoint intervals $\partial F =$ interval endpoints Field terms weight configurations by boundary placement.